AD-A110 545 RAND CORP SANTA MONICA CA F/G 12/1

ALGORITHMIC APPROXIMATION OF OPTIMAL VALUE DIFFERENTIAL STABILI--ETC(U)

AUG 81 W P HUTZLER

UNCLASSIFIED RAND/P-6659 NL

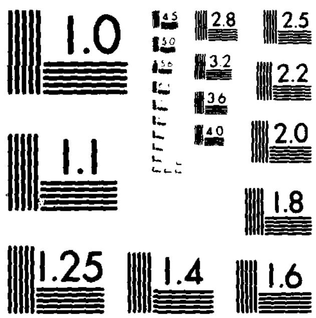
1 OF 1

END

DATE FILMED

8-82

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

2

12 27

ALGORITHMIC APPROXIMATION OF OPTIMAL VALUE DIFFERENTIAL STABILITY BOUNDS IN NONLINEAR PROGRAMMING

William P. Hutzler

August 1981

296600

P-6659

INTRODUCTION

The data needed to calculate the sensitivity to data perturbations of the solution and optimal value of a mathematical program are available as by-products of many solution algorithms. Fiacco ~~[8]~~ demonstrated this is developing a penalty function technique for approximating the parameter derivatives of a solution for quite general perturbed nonlinear programs. Armacost and Mylander ~~[3]~~ used this to advantage in making available the routine calculation of sensitivity information as part of a computer code for the Sequential Unconstrained Minimization Technique (SUMT). ~~[24]~~. Thus, based on the differentiable stability theory developed by Fiacco [8] and implemented by Armacost [1] and Armacost and Mylander [3], a number of stability measures, namely the parameter gradients of the optimal value function and the Karush-Kuhn-Tucker triple, can be represented and approximated using, for instance, only data needed for the SUMT solution technique itself.

Tight upper and lower bounds on the directional derivative limit quotients of the optimal value function for different forms of nonlinear programs with data perturbations have appeared in Gauvin and Dubeau [13], Fiacco and Hutzler [10], Fiacco [7], and Hutzler [17]. One form of these bounds is given by a pair of linear programs, each of whose feasible region is the set of Karush-Kuhn-Tucker multiplier vectors associated with a particular solution of the nonlinear program.

In this paper, we first consider equality constrained programs and determine conditions under which the optimal value directional derivative can be calculated using only first-order information about the problem functions. We also give conditions under which this directional derivative can be estimated using the iterates of any sequential solution technique.

Next we consider the more general pertur ogram containing inequality as well as equality constraints. Using mixed interior-exterior penalty function, we show that, when the parameter directional derivative of the optimal value function exists, it can be approximated or bounded above, depending on the nature of the solution generated by the penalty function algorithm. Moreover, we establish the existence of the parameter directional derivative of the mixed interior-exterior penalty function and obtain a representation of it.

NOTATION AND DEFINITIONS

In this paper we shall be concerned with mathematical programs of the general form:

$$\min_x f(x,\varepsilon) \qquad P(\varepsilon)$$

$$\text{s.t.} \quad g_i(x,\varepsilon) \geq 0 \ (i = 1,\ldots,m)$$

$$h_j(x,\varepsilon) = 0 \ (j = 1,\ldots,p),$$

where $x\varepsilon E^n$ is an n-dimensional vector of decision variables, ε is a parameter vector in E^k, and the functions f, g_i, and h_j, unless specified otherwise, are once continuously differentiable in x and ε. The m-vector whose components are $g_i(x,\varepsilon)$, $i = 1,\ldots,m$, and the p-vector whose components are $h_j(x,\varepsilon)$, $j = 1,\ldots,p$, will be denoted by $g(x,\varepsilon)$ and $h(x,\varepsilon)$, respectively.

Following usual conventions, all vectors in standard form, except gradient vectors, will be understood to be column vectors. The gradient, with respect to x, of a once differentiable real-valued function $f:E^n \times E^k \to E^1$ will be denoted $\nabla_x f(x,\varepsilon)$ and is taken to be the row vector $[\partial f(x,\varepsilon)/\partial x_1,\ldots,\partial f(x,\varepsilon)/\partial x_n]$. The transpose of a row or column vector y will be denoted y'. If $g(x,\varepsilon)$ is a vector-valued function, $g:E^n \times E^k \to E^m$, whose components $g_i(x,\varepsilon)$ are differentiable in x, then $\nabla_x g(x,\varepsilon)$ will denote the m x n Jacobian matrix of g whose ith row is given by $\nabla_x g_i(x,\varepsilon)$, $i = 1,\ldots,m$. The transpose of the Jacobian $\nabla_x g(x,\varepsilon)$ will be denoted $\nabla'_x g(x,\varepsilon)$. Differentiation with respect to the vector ε is denoted in a completely analogous fashion.

The feasible region of problem $P(\varepsilon)$ is the set of points $x\varepsilon E^n$ which satisfy the constraints and will be denoted $R(\varepsilon)$. Thus

$$R(\varepsilon) = \{x: g(x,\varepsilon) \geq 0, \ h(x,\varepsilon) = 0\}.$$

The set of solutions of $P(\varepsilon)$ will be denoted $S(\varepsilon)$ and is given by

$$S(\varepsilon) = \{x\varepsilon R(\varepsilon): f(x,\varepsilon) = f^*(\varepsilon)\},$$

where $f^*(\varepsilon)$, the optimal value function for $P(\varepsilon)$, is expressed as

$$f^*(\varepsilon) = \min \{f(x,\varepsilon): x \varepsilon R(\varepsilon)\}.$$

The Lagrangian for $P(\varepsilon)$ will be written

$$L(x,\mu,\omega,\varepsilon) = f(x,\varepsilon) - \sum_{i=1}^{m} \mu_i g_i(x,\varepsilon) + \sum_{j=1}^{p} \omega_j h_j(x,\varepsilon),$$

and the set of Karush-Kuhn-Tucker vectors [19,20] corresponding to the decision vector x will be given by

$$K(x,\varepsilon) = \{(\mu,\omega) \varepsilon E^m \times E^p: \nabla_x L(x,\mu,\omega,\varepsilon) = 0,$$

$$\mu_i \geq 0,\ \mu_i g_i(x,\varepsilon) = 0,\ i = 1,\dots,m\}.$$

Finally, writing a solution vector as a function of the parameter ε, the index set for inequality constraints which are binding at a solution $x(\varepsilon)$ will be denoted $B(\varepsilon) = \{i: g_i(x[\varepsilon],\varepsilon) = 0\}$.

The program $P(\varepsilon)$ specifies the general mathematical program with data perturbations. There are a number of forms of $P(\varepsilon)$ which are often used to represent problems with special structure. Those that we shall have occasion to address are the equality constrained program

$$\min f(x,\varepsilon) \qquad P_1(\varepsilon)$$

$$\text{s.t.} \quad h(x,\varepsilon) = 0,$$

and the right-hand-side problem

$$\min f(x) \qquad\qquad P_2(\varepsilon)$$

$$\text{s.t.} \quad g(x) \geq \varepsilon_1,$$

$$h(x) = \varepsilon_2,$$

where $\varepsilon = (\varepsilon_1, \varepsilon_2)$ is a vector in $E^m \times E^p$. The feasible regions and solution sets associated with $P_1(\varepsilon)$ and $P_2(\varepsilon)$ will, as for $P(\varepsilon)$, be denoted $R(\varepsilon)$ and $S(\varepsilon)$, respectively. The interpretation of this notation should be clear from the context in which it is applied. Other variations of $P(\varepsilon)$ will be explained as they arise in subsequent discussion.

Constraint qualifications in mathematical programming are regularity conditions which are generally imposed to ensure that the set of Karush-Kuhn-Tucker multipliers corresponding to an optimal solution is nonempty. One constraint qualification which we shall employ throughout this paper is the well-known Mangasarian-Fromovitz Constraint Qualification (MFCQ) which holds at a point $x \varepsilon R(\varepsilon)$ if:

i) there exists a vector $\tilde{y} \varepsilon E^n$ such that

$$\nabla_x g_i(x,\varepsilon)\ \tilde{y} > 0 \text{ for } i \varepsilon B(\varepsilon), \text{ and}$$

$$\nabla_x h_j(x,\varepsilon)\ \tilde{y} = 0 \text{ for } j = 1,\ldots,p; \text{ and}$$

ii) the gradients $\nabla_x h_j(x,\varepsilon)$, $j = 1,\ldots,p$, are linearly independent.

Another constraint qualification we shall have occasion to use is the linear independence condition, designated LI, which holds at a point $x \varepsilon R(\varepsilon)$ if the gradients $\{\nabla_x g_i(x,\varepsilon),\ i \varepsilon B(\varepsilon);\ \nabla_x h_j(x,\varepsilon),\ j = 1,\ldots,p\}$ are linearly independent. Clearly, for equality constrained programs, MFCQ and LI are equivalent regularity conditions.

Robinson [25] has shown the equivalence of MFCQ and a form of local stability of the solution set of a system of inequalities. Gauvin [12] has shown that MFCQ is both necessary and sufficient for the set $K(x,\varepsilon)$ to be nonempty, compact, and convex. In addition, Gauvin and Tolle [14] established that MFCQ is preserved under right-hand-side perturbations. That is, if MFCQ holds at a point $x^* \varepsilon S(\varepsilon^*)$ of $P_2(\varepsilon)$, and if $x_n \to x^*$ with $x_n \varepsilon S(\varepsilon_n)$ where $\varepsilon_n \to \varepsilon^*$, then, for n sufficiently large, MFCQ holds at x_n with $(\mu_n,\omega_n)\varepsilon K(x_n,\varepsilon)$ such that $(\mu_n,\omega_n) \to (\mu^*,\omega^*)$ for some $(\mu^*,\omega^*)\varepsilon K(x^*,\varepsilon^*)$.

We shall also have occasion to make use of the following compactness property for point-to-set maps.

<u>Definition 1</u>. A point-to-set mapping $\phi: X \to Y$ is said to be uniformly compact near $\bar{\varepsilon}$ of X if the closure of the set $\cup\phi(\varepsilon)$, where the union is over all ε in $N(\bar{\varepsilon})$, is compact for some neighborhood $N(\bar{\varepsilon})$ of $\bar{\varepsilon}$.

DIFFERENTIAL STABILITY OF THE OPTIMAL VALUE FUNCTION

Results relating to the differential stability of the optimal value function have generally been directed toward the existence of its directional derivative or bounds on its directional derivative limit quotients. Before describing these results and their relationship to results in this paper, we state several relevant definitions. Throughout the remainder of this work, we shall, without loss of generality, focus attention on the specific parameter value $\varepsilon = 0$ when considering the program $P(\varepsilon)$ and its variations.

<u>Definition 2</u>. For z, a unit vector in E^k, the directional derivative of $f^*(0)$ in the direction z is defined to be:

$$D_z f^*(0) = \lim_{\beta \to 0^+} \frac{f^*(\beta z) - f^*(0)}{\beta}, \tag{1}$$

provided that the limit exists.

If the limit in (1) does not exist, then attention is focused on the upper and lower directional derivative limit quotients of $f^*(0)$, which are respectively defined as follows:

$$D_z^+ f^*(0) = \limsup_{\beta \to 0^+} \frac{f^*(\beta z) - f^*(0)}{\beta}; \qquad (2)$$

$$D_z^- f^*(0) = \liminf_{\beta \to 0^+} \frac{f^*(\beta z) - f^*(0)}{\beta}. \qquad (3)$$

Clearly the limits in (2) and (3) always exist, though they may not be finite. Furthermore, $D_z f^*(0)$ exists if and only if these two limits are equal, in which case $D_z f^*(0) = D_z^+ f^*(0) = D_z^- f^*(0)$.

One of the earliest characterizations of the optimal value function of a mathematical program was provided by Danskin [5,6]. Addressing the problem minimize $f(x,\varepsilon)$ subject to $x \varepsilon R$, R some topological space, ε in E^k, Danskin derived conditions under which the directional derivative of f^* exists and also determined its representation. In particular, he showed that if R is nonempty and compact, and if f and its partial derivatives $\partial f/\partial \varepsilon_i$, $i = 1,\ldots,k$, are continuous at the point $\varepsilon = 0$, then the directional derivative of $f^*(0)$ exists and is given by

$$D_z f^*(0) = \min_{x \varepsilon S(0)} \nabla_\varepsilon f(x,0) z, \qquad (4)$$

where $S(0) = \{x\colon x \text{ minimizes } f(x,0) \text{ over } R\}$.

Although Danskin's result has been extended to other spaces and a variety of functional forms, its principal restriction remains that the feasible region, R, does not vary with the parameter ε. However, this situation does not prevent Danskin's result from being applied to problems such as $P(\varepsilon)$ since inequality and equality constraints can be "absorbed" into the objective function of a program by the use of an appropriate auxiliary function, e.g., Lagrangian or penalty function. Note also that it can be applied directly to the dual of a convex program with right-hand-side perturbations.

For the special case in which R is defined by inequalities of the form $g(x,\varepsilon) \geq 0$ and f and -g are convex and continuously differentiable on $X\varepsilon E^n$ for each fixed ε, Hogan [16] has given conditions under which $D_z f^*(0)$ exists and is finite for all $z \varepsilon E^k$. For this problem, the additional conditions ensure that S(0) is nonempty and bounded and that $g(\bar{x},0) > 0$ for some $\bar{x}\varepsilon X$. Under these conditions,

$$D_z f^*(0) = \min_{x\varepsilon S(0)} \max_{\mu\varepsilon K(x,0)} [\nabla_\varepsilon f(x,0) - \mu' \nabla_\varepsilon g(x,0)]z. \qquad (5)$$

For programs without equality constraints, Rockafellar [26] has shown that, under certain second-order conditions, the optimal value function of $P_2(\varepsilon)$ satisfies a stability of degree two, i.e., in a neighborhood of $\varepsilon = 0$, there exists a twice differentiable function $\phi: E^m \to E^1$ with $f^*(\varepsilon) \geq \phi(\varepsilon)$ and $f^*(0) = \phi(0)$. Under this stability property, bounds on the directional derivative of f^* can be derived. For convex programming problems of the form $P(\varepsilon)$, Gol'stein [15] has shown that a saddle point condition is satisfied by the directional derivative of f^*.

Recent investigations of Gauvin and Tolle [14] have focused on the stability of the optimal value function for programs of the form $P_2(\varepsilon)$, whose equality and inequality constraints are subject to right-hand-side perturbations. Assuming the Mangasarian-Formovitz constraint qualification and the uniform compactness of $R(\varepsilon)$, but without requiring the convexity of the problem functions or second-order conditions, Gauvin and Tolle showed that

$$D_z^+ f^*(0) \leq \inf_{x\varepsilon S(0)} \max_{(\mu,\omega)\varepsilon K(x,0)} \left(\sum_{i=1}^{m} \mu_i z_i - \sum_{j=1}^{p} \omega_j z_{m+j} \right) \qquad (6)$$

and

$$D_z^- f^*(0) \geq \inf_{x \in S(0)} \min_{(\mu,\omega) \in K(x,0)} \left(\sum_{i=1}^{m} \mu_i z_i - \sum_{j=1}^{p} \omega_j z_{m+j} \right). \quad (7)$$

It is interesting to note that in all of the works just mentioned the functional form that results in portraying the stability of the optimal value function can be expressed as an optimization of $\nabla_\varepsilon L(x,\mu,\omega,0)z$. This observation is pursued in [7,10,13,17] to show that the above results do indeed generalize to problems of the form $P(\varepsilon)$, where the parameter can appear anywhere in the objective function and in the inequality and equality constraints. The conditions imposed by Gauvin and Tolle [14] are shown to be sufficient to permit the extension of (6) and (7) to the more general program $P(\varepsilon)$. The next lemma states that extension in a form most suited to the needs of this paper.

Lemma 3. If, for $P(\varepsilon)$, $R(0)$ is nonempty and MFCQ holds at each $x \varepsilon S(0)$, then for any unit vector $z \varepsilon E^k$,

$$D_z^+ f^*(0) \leq \inf_{x \in S(0)} \max_{(\mu,\omega) \in K(x,0)} \nabla_\varepsilon L(x,\mu,\omega,0)\, z. \quad (8)$$

and if $R(\varepsilon)$ is uniformly compact for ε near $\varepsilon = 0$, then

$$D_z^- f^*(0) \geq \min_{(\mu,\omega) \in K(x^*,0)} \nabla_\varepsilon L(x^*,\mu,\omega,0)\, z \quad (9)$$

holds for some $x^* \varepsilon S(0)$.

Lempio and Maurer [21] have obtained similar bounds under analogous assumptions that are required to handle general perturbed infinite dimensional programs of the form minimize $f(x,\varepsilon)$, subject to $x \varepsilon R_1$ and $g(x,\varepsilon) \varepsilon R_2$, where R_1 and R_2 are arbitrary closed convex sets. Auslender [4] has also obtained these bounds for the right-hand-side perturbation problem $P(\varepsilon)$, extending the results of Gauvin and Tolle [14] by using a weaker form of the Mangasarian-Fromovitz constraint qualification. This allows him to replace the differentiability assumption on the

objective and inequality functions with the weaker requirement that they be locally Lipschitz.

Although we are focusing attention on programs for which the spaces involved are finite dimensional, we note that most of the sensitivity results mentioned above have been extended to infinite dimensional programs. For example, Maurer [22,23] has recently obtained a characterization of the subgradient of the extremal value function for the problem $P(\varepsilon)$, and has applied his results to a class of optimal control problems.

CALCULATING STABILITY BOUNDS FOR EQUALITY CONSTRAINED PROGRAMS

Consider the equality constrained program $P_1(\varepsilon)$. In [17] we have shown that when the feasible region is nonempty and uniformly compact near $\varepsilon = 0$, if the Jacobian matrix $\nabla h(x,0)$ has rank p for each $x \in S(0)$ then

$$D_z f^*(0) = \min_{x \in S(0)} \nabla_\varepsilon L(x,\omega[x,0],0)\, z, \tag{10}$$

for any unit vector $z \in E^k$. Here, because of the linear independence of the constraint gradients at each $x \in S(0)$, $\omega(x,0)$ is the unique multiplier vector associated with x.

Now, to calculate $D_z f^*(0)$ from (10), one must solve a typically nonlinear program. Two special cases in which this situation can be avoided are when S(0) is a singleton and when, for each unit vector $z \in E^k$, $D_z f^*(0) = -D_{-z} f^*(0)$. We shall now show that in the latter case, $D_z f^*(0)$ can be calculated explicitly.

We first note that, from (10)

$$D_z f^*(0) \le \nabla_\varepsilon L(x,\omega[x,0],0)\, z$$

for any $x \in S(0)$, so, substituting $-z$ for z we obtain

$$-D_{-z} f^*(0) \ge \nabla_\varepsilon L(x,\omega[x,0],0)\, z.$$

Thus, if, in addition to the hypotheses which guarantee (10), we have that $D_z f^*(0) = -D_{-z} f^*(0)$, then

$$D_z f^*(0) = \nabla_\varepsilon L(x,\omega[x,0],0)\ z \qquad (11)$$

for any $x \varepsilon S(0)$. Again, $\omega(x,0)$ is the unique multiplier vector associated with x. All that remains to determine $D_z f^*(0)$ explicitly is to calculate $\omega(x,0)$ which, by the first-order necessary condition, must satisfy

$$-\omega' \nabla_x h(x,0) = \nabla_x f(x,0). \qquad (12)$$

Under the assumption that $\nabla_x h(x,0)$ has full row rank for each $x \varepsilon S(0)$, we see that (12) has a unique solution given by

$$\omega(x,0) = -\nabla_x h(x,0)\ (\nabla_x' h[x,0]\ \nabla_x h[x,0])^{-1}\ \nabla_x' f(x,0). \qquad (13)$$

Substituting this expression into (11) we see that $D_z f^*(0)$ is given by

$$\nabla_\varepsilon f(x,0)\ z - \nabla_x f(x,0)\ (\nabla_x' h[x,0]\ \nabla_x h[x,0])^{-1}\ \nabla_x' h(x,0)\ \nabla_\varepsilon h(x,0) z \quad (14)$$

for any $x \varepsilon S(0)$, under the condition that $D_z f^*(0) = -D_{-z} f^*(0)$.

If, on the other hand, S(0) is a singleton, i.e., if x^* is the unique solution of $P_1(0)$, then, from (10) it follows that (14) is an exact representation of $D_z f^*(0)$ and holds without the added hypothesis that $D_z f^*(0) = -D_{-z} f^*(0)$.

Suppose now that we have a sequence of points $\{x_k\}$, given by some algorithm for solving $P_1(\varepsilon)$, which converges to $x^* \varepsilon S(0)$. Since our problem functions are assumed to be once continuously differentiable, the expression for $D_z f^*(0)$ given in (14) is a continuous function of x. Thus for k sufficiently large, (14), when evaluated at x_k, provides an estimate of $D_z f^*(0)$ if x^* is unique. If x^* is not unique, then (14), when evaluated at x_k for k sufficiently large, provides an

estimate of $D_z f^*(0)$ whenever $D_z f^*(0) = -D_{-z} f^*(0)$, and an upper bound on $D_z f^*(0)$ whenever $D_z f^*(0) \neq -D_{-z} f^*(0)$.

PENALTY FUNCTION DIRECTIONAL DERIVATIVES

Before developing similar results for the more general program, $P(\varepsilon)$, we briefly review the fundamental convergence properties of penalty functions. These properties are employed below to approximate $D_z f^*(\varepsilon)$. A detailed discussion of the properties of penalty functions can be found in Fiacco and McCormick [11].

A general form of the mixed interior-exterior penalty function for $P(\varepsilon)$ can be written

$$V(x,\varepsilon,r_k,t_k) = f(x,\varepsilon) + s(r_k)\ I(x,\varepsilon) + p(t_k)\ G(x,\varepsilon), \qquad (15)$$

where $I = I(g[x,\varepsilon])$ and $G = G(h[x,\varepsilon])$, and assuming the following:

i) I is a continuous function on $R''(\varepsilon) = \{x\colon g(x,\varepsilon) > 0\}$,

ii) if $\{x_k\} \subseteq R''(\varepsilon)$ with $x_k \to x^*$ and $g_i(x^*,\varepsilon) = 0$ for at least one $i = 1,\ldots,m$, then $\lim_{k \to \infty} I(x_k \varepsilon) = +\infty$;

iii) G is continuous with $G(x,\varepsilon) = 0$ if $h(x,\varepsilon) > 0$ and $G(x,\varepsilon) > 0$ otherwise;

iv) $r_m > r_n > 0$ implies $s(r_m) > s(r_n) > 0$;

v) $\lim_{k \to \infty} r_k = 0$ implies $\lim_{k \to \infty} s(r_k) = 0$;

vi) $0 < t_m < t_n$ implies $0 < p(t_m) < p(t_n)$; and

vii) if $k \to +\infty$ monotonically, then $\lim_{k \to \infty} p(t_k) = \infty$.

It is by now well known that the sequential unconstrained minimization of the penalty function (15) can serve as a means of approximating local solutions of $P(\varepsilon)$. The conditions under which this is so are contained in Lemma 4, which is due to Fiacco and McCormick [11]. Let $H(\varepsilon) = \{x\colon h(x,\varepsilon) = 0\}$.

Lemma 4. If,

a) f, g_i $(i = 1,\ldots,m)$, and h_j $(j = 1,\ldots,p)$ are continuous;

b) $I(x,\varepsilon)$, $G(x,\varepsilon)$, $s(r)$, and $p(t)$ satisfy (i)-(vii) above;

c) a set, $S'(\varepsilon)$ of local minima of $P(\varepsilon)$ is a nonempty isolated compact set;

d) $S(\varepsilon) \cap cl(R''[\varepsilon] \cap H[\varepsilon]^o) \neq \Phi$; and

e) $r_k \to 0$ and $t_k \to +\infty$, each monotonically with $t_k > 0$,

then

i) there exists a compact set $S''(\varepsilon)$ with $S'(\varepsilon) \subset S''(\varepsilon)^o$, and for k sufficiently large, the unconstrained minima of $V(x,\varepsilon,r_k,t_k)$ in $R'' \cap S''(\varepsilon)^o$ exist and every limit point of any sequence $\{x_k\}$ of minimizing points is in $S'(\varepsilon)$;

ii) $\lim_{k \to \infty} s(r_k)\, I(x_k,\varepsilon) = 0$;

iii) $\lim_{k \to \infty} p(t_k)\, G(x_k,\varepsilon) = 0$; and

iv) $\lim_{k \to \infty} V(x_k,\varepsilon,r_k,t_k) = f^*(\varepsilon)$.

Now consider the mathematical program $P(\varepsilon)$, which contains inequality constraints as well as equality constraints. Recall that for such a program, the general form of the mixed interior-exterior penalty function is

$$V(x,\varepsilon,r_k,t_k) = f(x,\varepsilon) + s(r_k)\, I(x,\varepsilon) + p(t_k)\, G(x,\varepsilon), \qquad (16)$$

and that one method for solving $P(\varepsilon)$ is to solve the sequence of unconstrained minimization problems

$$\min_x V(x,\varepsilon,r_k,t_k). \qquad P_k(\varepsilon)$$

Denoting by $V_k^*(\varepsilon)$ the optimal value function of the program $P_k(\varepsilon)$, we now establish the existence of the directional derivative $D_z V_k^*(\varepsilon)$. In doing so, we assume, in addition to the hypotheses of Lemma 4, that the functions $I(x,\varepsilon)$ and $G(x,\varepsilon)$ are once continuously differentiable.

Under our hypotheses, $V(x,\varepsilon,r_k,t_k)$ is continuously differentiable, $\nabla_\varepsilon V$ exists, and V and $\nabla_\varepsilon V$ are continuous in x and ε. In addition, by Lemma 4, the unconstrained minima of $P_k(0)$ lie in $R''(0) \cap S''(0)^o$ for k sufficiently large, where $S''(0)$ is compact. Thus, taking $X = S''(0)$ in Danskin's theorem, it follows that $D_z V_k^*(0)$ exists for all directions $z \varepsilon E^k$ and all indices k sufficiently large. Furthermore,

$$D_z V_k^*(0) = \min_{x \varepsilon S_k(0)} \nabla_\varepsilon V(x,0,r_k,t_k)\ z, \tag{17}$$

where $S_k(0) = \{x \varepsilon S''(0):\ V_k^*(0) \geq V(x,0,r_k,t_k)\}$. In particular, if x_k^* is the unique minimizer of $V_k(0)$, then

$$D_z V_k^*(0) = \nabla_\varepsilon V(x_k^*,0,r_k,t_k)\ z.$$

As another consequence of (17) we note that, if k is sufficiently large and $x_k = x_k(r_k,t_k,0)$ is a minimizer of $P_k(0)$, i.e., $x_k \varepsilon S_k(0)$, then, for any unit vector $z \varepsilon E^k$,

$$D_z V_k^*(0) \leq \nabla_\varepsilon V(x_k,0,r_k,t_k)\ z. \tag{18}$$

Several observations now follow. First, from (18) we see that

$$-D_{-z} V_k^*(0) \geq \nabla_\varepsilon V(x_k,0,r_k,t_k)\ z, \tag{19}$$

so that, combining (18) and (19),

$$D_z V_k^*(0) \leq \nabla_\varepsilon V(x_k,0,r_k,t_k)\ z \leq -D_{-z} v_k^*(0).$$

Thus, if $D_z V_k^*(0) = -D_{-z} V_k^*(0)$ for some unit vector $z^* \epsilon E^k$, then

$$D_{z^*} V_k^*(0) = \nabla_\epsilon V(x_k, 0, r_k, t_k)\; z^* \tag{20}$$

for any minimizer x_k of $P_k(0)$. If $D_z f^*(0) = -D_{-z} f^*(0)$ for all unit vectors $z \epsilon E^k$, then (20) holds for all direction vectors z and this provides a necessary condition for the existence of $\nabla_\epsilon V_k^*(0)$.

If we now take $V(x,0,r_k,t_k)$ to be the logarithmic-quadratic penalty function

$$W(x,0,r_k) = f(x,0) - r_k \sum_{i=1}^{m} \ln\, g_i(x,0) + \frac{1}{2r_k} \sum_{j=1}^{p} h_j^2(x,0), \tag{21}$$

then, from (17), for k sufficiently large,

$$D_z W_k^*(0) = \min_{x \epsilon S_k(0)} \nabla_\epsilon W(x,0,r_k)\; z \tag{22}$$

for any direction $z \epsilon E^k$.

Now, choosing any sequence of minimizers $\{x_k\}$ with $x_k \epsilon S_k(0)$, we have, by Lemma 4, that $x_k \to x^*$ where x^* is a minimizer of P(0). Thus, by the continuity of $\nabla_\epsilon W$, $\nabla_\epsilon W(x_k,0,r_k) \to \nabla_\epsilon L(x^*,\mu^*,\omega^*,0)$ as $k \to \infty$, where the multipliers μ^* and ω^* are given by

$$\mu_i^* = \lim_{k \to \infty} \frac{r_k}{g_i(x_k,0)}, \qquad i = 1,\ldots,m,$$

$$\omega_j^* = \lim_{k \to \infty} \frac{h_j(x_k,0)}{r_k}, \qquad j = 1,\ldots,p.$$

Clearly, if MFCQ holds at each point of S(0), then (8) and (9) apply. Thus $\nabla_\epsilon W(x_k,0,r_k)\; z$, which converges to $\nabla_\epsilon L(x^*,\mu^*,\omega^*,0)\; z$,

in general converges to a value between the bounds expressed in (8) and (9).

In the particular instance in which $D_z f^*(0)$ exists and is given by

$$D_z f^*(0) = \inf_{x \in S(0)} \nabla_\varepsilon L(x,\mu[x,0],\omega[x,0],0)\ z,$$

as for instance when each $x \in S(0)$ possesses a unique multiplier vector, we may conclude that

$$D_z f^*(0) \leq \nabla_\varepsilon L(x^*,\mu[x^*,0],\omega[x^*,0],0)\ z \qquad (23)$$

for any $x^* \in S(0)$. Here, for k sufficiently large, $\nabla_\varepsilon W(x_k,0,r_k)\ z$ approximates $D_z f^*(0)$ if equality holds in (23), as when P(0) has a unique solution, or $\nabla_\varepsilon W(x_k,0,r_k)\ z$ is a strict upper bound on $D_z f^*(0)$ if equality does not hold in (23).

__Example__

To demonstrate a number of the concepts just discussed, consider the problem

$$\min\ \varepsilon x_1 \qquad\qquad P(\varepsilon)$$

$$\text{s.t.}\quad g(x,\varepsilon) = -(x_1 - \varepsilon)^2 - (x_2 + 2)^2 + 4 \geq 0$$

$$h(x,\varepsilon) = -x_1 + x_2 + \varepsilon = 0$$

The solution of this program is easily determined to be $x_1^* = x_2^* + \varepsilon$ with

$$x_2^* = \begin{cases} 0 & \varepsilon > 0 \\ 2 & \varepsilon < 0, \text{ and if } \varepsilon = 0,\ x_2^* \text{ can be any value} \\ & \text{in the interval } [0,2]. \end{cases} \tag{24}$$

The optimal value function is

$$f^*(\varepsilon) = \begin{cases} \varepsilon^2 & \varepsilon \geq 0 \\ \varepsilon^2 + 2\varepsilon & \varepsilon < 0 \end{cases} . \tag{25}$$

We see that f^* is a continuous function of the parameter ε, but it is not differentiable at $\varepsilon = 0$. This function does, however, have directional derivatives at $\varepsilon = 0$ which are given by

$$D_z f^*(0) = \begin{cases} 0 & z = 1 \\ -2 & z = -1 \end{cases} . \tag{26}$$

When $\varepsilon \neq 0$, the derivative of $f^*(\varepsilon)$ is

$$\frac{df^*}{d\varepsilon} = \begin{cases} 2\varepsilon & \varepsilon \geq 0 \\ 2\varepsilon + 2 & \varepsilon < 0 \end{cases} . \tag{27}$$

The logarithmic-quadratic penalty function for $P(\varepsilon)$ is

$$W(x,\varepsilon,r) = \varepsilon x_1 - r\, ln[-(x_1 - \varepsilon)^2 - (x_2 + 2)^2 + 4] + \frac{1}{2r} [-x_1 + x_2 + \varepsilon].$$

Choosing a sequence $\{r_k\}$ of positive real numbers with $r_k \to 0$ monotonically, we have solved the sequence of minimization problems

$$\min_x W(x,\varepsilon,r_k) \qquad P_k(\varepsilon)$$

using the Sequential Unconstrained Minimization Technique (SUMT) developed by Fiacco and McCormick [11] and implemented by Mylander, Holmes, and McCormick [24]. The results of the minimization of the sequence of problems $P_k(\varepsilon)$ for $\varepsilon = -1$ are shown in Table 1.

The sequence of points, $x_k = x_k(r_k, \varepsilon)$, generated by SUMT can be used to bound the directional derivative $D_z f^*(\varepsilon)$ as in (23), since $x_k \varepsilon S_k(\varepsilon)$ and, as we shall see, almost each $x \varepsilon S(\varepsilon)$ possesses a unique multiplier. To this end, consider the gradient of the Lagrangian for $P(\varepsilon)$, i.e.,

$$\nabla L(x, \mu, \omega, \varepsilon) = (\varepsilon + 2\mu[x_1 - \varepsilon] - \omega,\ 2\mu[x_2 - 2] + \omega). \qquad (28)$$

A first-order necessary condition that $x^* = (x_1^*, x_2^*)$ be a local minimum of $P(\varepsilon)$ is that there exist vectors $\mu(x^*, \varepsilon)$ and $\omega(x^*, \varepsilon)$ such that $\nabla L(x^*, \mu(x^*, \varepsilon), \omega(x^*, \varepsilon), \varepsilon) = 0$. From (28) we see that $\mu^* = \mu(x^*, \varepsilon)$ and $\omega^* = \omega(x^*, \varepsilon)$ must satisfy the system of equations

$$\begin{bmatrix} 2(x_1^* - \varepsilon) & -1 \\ 2(x_2^* - 2) & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \omega \end{bmatrix} = \begin{bmatrix} -\varepsilon \\ 0 \end{bmatrix}. \qquad (29)$$

Recalling that at a solution of $P(\varepsilon)$, $x_1^* = x_2^* + \varepsilon$, it is easily shown that the system in (29) has a unique solution for any value of $\varepsilon \neq 0$. Furthermore, if $\varepsilon = 0$, the system has a unique solution provided $x_2^* \neq 1$. In those instances when the multiplier is unique, it is given by

$$\begin{bmatrix} \mu^* \\ \omega^* \end{bmatrix} = \frac{1}{4(x_2^* - 1)} \begin{bmatrix} -\varepsilon \\ -2\varepsilon(x_2^* - 2) \end{bmatrix}.$$

Thus, for the particular situation in which $\varepsilon = -1$, we can bound $D_z f^*(-1)$ as in (23). As indicated in the discussion following (23),

Table 1

THE SOLUTION OF P(-1) USING SUMT

k	r_k	$x_k^1(r_k,-1)$	$x_k^2(r_k,-1)$	g(x,-1)	h(x,-1)	$W_k^*(-1)$
1	100	-.9791	1.9993	3.9996	1.9784	.9791
2	10	-.7381	1.9343	3.9271	1.6724	.7381
3	1	.3774	1.5895	1.9342	.2121	-.3774
4	.1	.9049	1.9075	.3726	.0026	-.9049
5	.01	.9901	1.9901	.0396	.0000	-.9901
6	.001	.9990	1.9990	.0040	.0000	-.9990
7	.0001	.9999	1.9999	.0004	.0000	-.9999
8	.00001	1.0	2.0	.0000	.0000	-1.0

we can approximate the bounding term algorithmically by $\nabla_\varepsilon W(x_k,-1,r_k)\ z$, tor k sufficiently large. This is demonstrated using the figures provided in Table 1 and the formula for $\nabla_\varepsilon W(x_k,-1,r_k)$ which is

$$\nabla_\varepsilon W(x_k,-1,r_k) = x_k^1 - r\,\frac{2(x_k^1+1)}{g(x_k^1,-1)} + \frac{1}{r}\,h(x_k^2,-1).$$

The results of this sequence of calculations are shown in Table 2. As we can see, the sequence of values $\nabla_\varepsilon W(x_k,-1,r_k)$ is approximately zero for $k = 8$. Thus we may conclude, from the discussion following (23), that $D_z f^*(0) \leq 0$ for each $z \varepsilon \{-1,1\}$. This agrees with the exact calculation of $D_z f^*(0)$ implied by (27) since f^* is actually differentiable at $\varepsilon = -1$.

Unfortunately this approach cannot be used to bound the directional derivative of f^* when $\varepsilon = 0$. The reason is that not all solutions of P(0) have a unique multiplier vector. We have seen from (29) that when $\varepsilon = 0$, the solution $x^* = [1,1]$ has multiple, in fact infinitely many, multipliers associated with it.

RELATED RESULTS

The sensitivity of the optimal value function of a mathematical program has been the subject of a significant amount of research in recent years. Conditions ranging from the semi-continuity of problem functions to the assumption of strong second-order conditions and strict complementary slackness have resulted in a number of static measures of the local behavior of the optimal value function. The reader interested in a comprehensive review of these results, as well as other results in general mathematical programming sensitivity analysis, is referred to the survey by Fiacco and Hutzler [9].

Among the most recent of these results are those of Armacost and Fiacco [2], which show that if, at a local solution of $P(\varepsilon)$, the second-order sufficiency condition holds along with the linear independence of the binding constraint gradients and the strict complementary slackness condition, then the optimal value function,

Table 2

THE CALCULATION OF A BOUND ON $D_z f^*(0)$

	k	1	2	3	4	5	6	7	8
$\nabla_\varepsilon Wz$	z = 1	-2.0024	-1.9048	-.8347	-.1201	.4885	-.0026	.0056	.0055
	z = -1	2.0024	1.9048	.8347	.1201	-.4885	.0026	-.0058	-.0055

$f^*(\varepsilon)$, is twice continuously differentiable in ε. In addition, they obtained the form of f^* and its first and second derivatives in terms of the $P(\varepsilon)$ Lagrangian, its gradient, and its Hessian, respectively.

In [18], Jittorntrum showed that the strict complementary slackness condition is not essential to the differentiability of f^*, which results from the second-order sufficiency and linear independence conditions. Note that under these weakened conditions, however, that $\nabla_\varepsilon f^*$ need not be continuous and that higher order derivatives do not necessarily exist.

As we mentioned earlier, the differential stability of $f^*(\varepsilon)$ has been investigated by Gauvin and Tolle [14] for right-hand-side perturbations, and by Fiacco and Hutzler [10], Fiacco [7], Hutzler [17], and Gauvin and Dubeau [13], for more general programs of the form $P(\varepsilon)$. Each of these works produced bounds on the directional derivative limit quotients of f^*. Moreover, problems possessing special structure, e.g., convexity, are known to have optimal value functions with an at least theoretically computable directional derivative. Danskin [5,6] provided the classic result in this area in addressing unconstrained problems. His result readily applies to constrained problems since inequality and equality constraints can be "absorbed" into the objective function of a program by employing an appropriate auxiliary function (Lagrangian, penalty function, etc.). Gol'stein [15] and Hogan [16], each using different assumptions, demonstrated the existence of $D_z f^*$ for convex programs. Hogan used the form of that derivative to develop a convergent algorithm for solving decomposable convex programs. Aside from Hogan's application of $D_z f^*$, it does not appear that these results have been applied to the development or enhancement of algorithms for solving mathematical programs.

REFERENCES

1. Armacost, R. L., *Sensitivity Analysis in Parametric Nonlinear Programming*, Doctoral Dissertation, School of Engineering and Applied Science, The George Washington University, May 1976.

2. Armacost, R. L., and A. V. Fiacco, "Sensitivity Analysis for Parametric Nonlinear Programming Using Penalty Methods," in W. W. White (ed.), *Computers and Mathematical Programming*, National Bureau of Standards Special Publication 502, February 1978, pp. 261-269.

3. Armacost, R. L., and W. C. Mylander, "A Guide to the SUMT - Version 4 Computer Subroutine for Implementing Sensitivity Analysis in Nonlinear Programming," Technical Paper T-287, Institute for Management Science and Engineering, The George Washington University, July 1973.

4. Auslender, A., "Differential Stability in Non Convex and Non Differentiable Programming," *Mathematical Programming Studies*, Vol. 10, 1979, pp. 29-41.

5. Danskin, J. M., *The Theory of Max-Min*, Springer - Verlag Inc., New York, 1967.

6. -----, "The Theory of Max-Min with Applications," *SIAM Journal of Applied Mathematics*, Vol. 14, No. 4, 1966, pp. 651-665.

7. Fiacco, A. V., "Optimal Value Differential Stability for General Mathematical Programs," submitted to *Mathematical Programming*.

8. -----, "Sensitivity Analysis for Nonlinear Programming Using Penalty Methods," *Mathematical Programming*, Vol. 10, No. 3, 1976, pp. 287-311.

9. Fiacco, A. V., and W. P. Hutzler, "Basic Results in the Development of Sensitivity and Stability Analysis in Constrained Mathematical Programming," to appear in special issue of *Computers and Operations Research*, entitled "Mathematical Programming with Parameter and Multi-Level Constraints."

10. -----, "Extensions of Gauvin-Tolle Optimal Value Differential Stability Results to General Mathematical Programs," Technical Paper T-393, Institute for Management Science and Engineering, The George Washington University, April 1979.

11. Fiacco, A. V., and G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley and Sons, Inc., New York, 1968.

12. Gauvin, J., "A Necessary and Sufficient Regularity Condition to Have Bounded Multipliers in Nonconvex Programming," *Mathematical Programming*, Vol. 12, 1977, pp. 136-138.

13. Gauvin, J., and F. Dubeau, "Differential Properties of the Marginal Function in Mathematical Programming," Ecole Polytechnique de Montreal, June 1979.

14. Gauvin, J., and J. W. Tolle, "Differential Stability in Nonlinear Programming," *SIAM Journal of Control and Optimization*, Vol. 15, No. 2, 1977, pp. 294-311.

15. Gol'stein, E. G., *Theory of Convex Programming*, Translations of Mathematical Monographs, Vol. 36, American Mathematical Society, Providence, RI, 1972.

16. Hogan, W., "Directional Derivatives for Extremal Value Functions with Applications to the Completely Convex Case," *Operations Research*, Vol. 21, No. 1, 1973, pp. 188-209.

17. Hutzler, W. P., *Differential Stability of the Optimal Value Function in Constrained Nonlinear Programming*, Doctoral Dissertation, School of Engineering and Applied Science, The George Washington University, November 1980.

18. Jittorntrum, K., *Sequential Algorithms in Nonlinear Programming*, Doctoral Dissertation, Australian National University, July 1978.

19. Karush, W., *Minima of Functions of Several Variables with Inequalities as Side Conditions*, Master's Thesis, University of Chicago, 1939.

20. Kuhn, H. W., and A. W. Tucker, "Nonlinear Programming," in J. Neyman (ed.), *Proceedings Second Berkeley Symposium in Mathematical Statistics and Probability*, University of California Press, Berkeley, CA., 1951, pp. 481-492.

21. Lempio, F., and H. Maurer, "Differential Stability in Infinite-Dimensional Nonlinear Programs" to appear in *Applied Mathematics and Optimization*.

22. Maurer, H., "A Sensitivity Result for Infinite Nonlinear Programming Problems; Part I: Theory," Mathematisches Institut der Universität Wurzburg, Wurzburg, W. Germany, January 1977.

23. -----, "A Sensitivity Result for Infinite Nonlinear Programming Problems; Part II: Applications to Optimal Control Problems," Mathematisches Institut der Universität Wurzburg, Wurzburg, W. Germany, January 1977.

24. Mylander, W. C., R. L. Holmes, and G. P. McCormick, *A Guide to SUMT-Version 4*, RAC-P-63, Research Analysis Corporation, McLean, Virginia, 1971.

25. Robinson, S. M., "Stability Theory for Systems of Inequalities, Part I: Linear Systems," *SIAM Journal on Numerical Analysis*, Vol. 12, 1975, pp. 754-769.

26. Rockafellar, R. T., "Augmented Lagrange Multiplier Functions and Duality in Nonconvex Programming," *SIAM Journal on Control*, Vol. 12, No. 2, 1974, pp. 268-285.

FILMED

3-8